



Applicants' undersigned attorney may be reached in our New York office by telephone at (212) 218-2100. All correspondence should continue to be directed to our address given below.

Respectfully submitted,

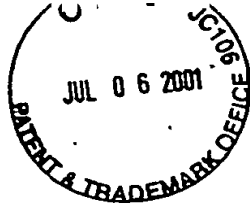
RECEIVED  
JUL 09 2001  
Technology Center 2600

Z. P. Diana  
Attorney for Applicants

Registration No. 78,496

FITZPATRICK, CELLA, HARPER & SCINTO  
30 Rockefeller Plaza  
New York, New York 10112-3801  
Facsimile: (212) 218-2200

NY\_MAIN 182271 v 1



09/817,345  
GAM: 2641

(translation of the front page of the priority document of  
Japanese Patent Application No. 2000-099418)

PATENT OFFICE  
JAPANESE GOVERNMENT

This is to certify that the annexed is a true copy of the  
following application as filed with this Office.

Date of Application: March 31, 2000

Application Number : Patent Application 2000-099418

Applicant(s) : Canon Kabushiki Kaisha

April 20, 2001

Commissioner,  
Patent Office

Kouzo OIKAWA

Certification Number 2001-3033155

CERTIFIED COPY OF  
PRIORITY DOCUMENT

CFM 410-00  
09/817,345  
GR4: 2641



日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office

出 願 年 月 日

Date of Application:

2000年 3月31日

出 願 番 号

Application Number:

特願2000-099418

出 願 人

Applicant(s):

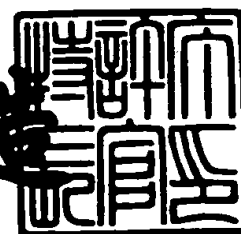
キヤノン株式会社

RECEIVED  
JUL 09 2001  
Technology Center 2600

2001年 4月20日

特許庁長官  
Commissioner,  
Japan Patent Office

及 川 耕 造



【書類名】 特許願

【整理番号】 4149142

【提出日】 平成12年 3月31日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/16

【発明の名称】 データ処理装置及び方法、ブラウザシステム、ブラウザ装置、記録媒体

【請求項の数】 28

【発明者】

    【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社  
社内

    【氏名】 伊藤 史朗

【発明者】

    【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社  
社内

    【氏名】 藤井 憲一

【発明者】

    【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社  
社内

    【氏名】 上田 隆也

【発明者】

    【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社  
社内

    【氏名】 池田 裕治

【特許出願人】

    【識別番号】 000001007

    【氏名又は名称】 キヤノン株式会社

【代理人】

    【識別番号】 100076428

【弁理士】

【氏名又は名称】 大塚 康德

【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100101306

【弁理士】

【氏名又は名称】 丸山 幸雄

【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100115071

【弁理士】

【氏名又は名称】 大塚 康弘

【電話番号】 03-5276-3241

【手数料の表示】

【予納台帳番号】 003458

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0001010

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 データ処理装置及び方法、ブラウザシステム、ブラウザ装置、記録媒体

【特許請求の範囲】

【請求項1】 ネットワーク上で提供されるデータの内容を、ブラウザ装置に対して音声データとして提供するためのデータ処理装置であって、

前記データに基づいて、その一部又は全部の内容を示す音声データを作成する手段と、

作成した音声データを格納する手段と、

前記データに、前記音声データを格納した場所を示す識別子を付加したデータを作成する手段と、

前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、  
を備えたことを特徴とするデータ処理装置。

【請求項2】 ネットワーク上で提供されるデータに対して、ブラウザ装置から音声により返答するためのデータ処理装置であって、

前記データの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する手段と、

前記データに、前記ブラウザ装置が音声データにより前記返答を行う場合の受付先を示す識別子を付加したデータを作成する手段と、

前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、  
を備えたことを特徴とするデータ処理装置。

【請求項3】 前記返答に係る音声データが、前記ブラウザ装置から前記受付先へ供給された場合に、前記音声データに対して、音声認識を行う認識手段を備えたことを特徴とする請求項2に記載のデータ処理装置。

【請求項4】 前記認識手段の結果に基づいて、前記返答を受け付ける前記ネットワーク上のサーバに適応した形式の返答データを作成する手段と、

前記返答データを前記サーバに提供する手段と、  
を備えたことを特徴とする請求項3に記載のデータ処理装置。

【請求項5】 前記返答が、予め定められた複数の項目から選択すべきもの

である場合に、

各々の前記項目に係る音声データを認識するための認識文法を作成する手段と

前記認識文法に基づいて、前記ブラウザ装置からの前記返答に係る音声データが、いずれの前記項目に相当するかを判定する手段と、

各々の前記項目に対応して前記返答を受け付ける前記ネットワーク上のサーバに適応した形式の返答データを作成する手段と、

前記返答データを前記サーバに提供する手段と、  
を備えたことを特徴とする請求項2に記載のデータ処理装置。

【請求項6】 前記返答データは、前記識別子を付加したデータを前記ブラウザ装置に対して提供する前に作成されることを特徴とする請求項5に記載のデータ処理装置。

【請求項7】 ブラウザ装置と、ネットワークを介して前記ブラウザ装置にデータを提供するサーバと、前記サーバにより提供されるデータの内容を、ブラウザ装置に対して音声データとして提供するためのデータ処理装置と、を備えたブラウザシステムであって、

前記データ処理装置は、

前記サーバが提供する前記データに基づいて、その一部又は全部の内容を示す音声データを作成する手段と、

作成した音声データを格納する手段と、

前記データに、前記音声データを格納した場所を示す識別子を付加したデータを作成する手段と、

前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、  
を備え、

前記ブラウザ装置は、前記識別子に示された場所から前記音声データを取得し、該音声データに係る音声を出力する手段を備えたことを特徴とするブラウザシステム。

【請求項8】 ブラウザ装置と、ネットワークを介して前記ブラウザ装置にデータを提供するサーバと、前記サーバにより提供されるデータに対して、ブラ



ウザ装置から音声により返答するためのデータ処理装置と、を備えたブラウザシステムであって、

前記データ処理装置が、

前記データの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する手段と、

前記データに、前記ブラウザ装置が音声データにより前記返答を行う場合の受付先を示す識別子を付加したデータを作成する手段と、

前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、

前記返答に係る音声データが、前記ブラウザ装置から前記受付先へ供給された場合に、前記音声データに対して、音声認識を行う認識手段と、

前記認識手段の結果に基づいて、前記返答を受け付ける前記サーバに適應した形式の返答データを作成する手段と、

前記返答データを前記サーバに提供する手段と、を備え、

前記ブラウザ装置が、

音声を入力する手段と、

入力された音声に基づいて音声データを作成する手段と、

作成した音声データを前記識別子が示す受付先へ供給する手段と、

を備えたことを特徴とするブラウザシステム。

【請求項9】 ブラウザ装置と、ネットワークを介して前記ブラウザ装置にデータを提供するサーバと、前記サーバにより提供されるデータの内容を、ブラウザ装置に対して音声データとして提供すると共に、前記サーバにより提供されるデータに対して、前記ブラウザ装置から音声により返答するためのデータ処理装置と、を備えたブラウザシステムであって、

前記データ処理装置は、

前記サーバが提供する前記データに基づいて、その一部又は全部の内容を示す音声データを作成する手段と、

作成した音声データを格納する手段と、

前記データに、前記音声データを格納した場所を示す第1の識別子を付加したデータを作成する手段と、

前記第1の識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、

前記データの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する手段と、

前記データに、前記ブラウザ装置が音声データにより前記返答を行う場合の受付先を示す第2の識別子を付加したデータを作成する手段と、

前記第2の識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、

前記返答に係る音声データが、前記ブラウザ装置から前記受付先へ供給された場合に、前記音声データに対して、音声認識を行う認識手段と、

前記認識手段の結果に基づいて、前記返答を受け付ける前記サーバに適應した形式の返答データを作成する手段と、

前記返答データを前記サーバに提供する手段と、を備え、

前記ブラウザ装置は、

前記第1の識別子に示された場所から前記音声データを取得し、該音声データに係る音声を出力する手段と、

音声を入力する手段と、

入力された音声に基づいて音声データを作成する手段と、

作成した音声データを前記第2の識別子が示す受付先へ供給する手段と、を備えたことを特徴とするブラウザシステム。

【請求項10】 ネットワーク上で提供されるデータの内容を、ブラウザ装置に対して音声データとして提供するためのデータ処理方法であって、

前記データに基づいて、その一部又は全部の内容を示す音声データを作成する工程と、

作成した音声データを格納する工程と、

前記データに、前記音声データを格納した場所を示す識別子を付加したデータを作成する工程と、

前記識別子を付加したデータを前記ブラウザ装置に対して提供する工程と、を含むことを特徴とするデータ処理方法。

【請求項11】 ネットワーク上で提供されるデータに対して、ブラウザ装置から音声により返答するためのデータ処理方法であって、

前記データの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する工程と、

前記データに、前記ブラウザ装置が音声データにより前記返答を行う場合の受付先を示す識別子を付加したデータを作成する工程と、

前記識別子を付加したデータを前記ブラウザ装置に対して提供する工程と、を含むことを特徴とするデータ処理方法。

【請求項12】 前記返答に係る音声データが、前記ブラウザ装置から前記受付先へ供給された場合に、前記音声データに対して、音声認識を行う認識工程を含むことを特徴とする請求項11に記載のデータ処理方法。

【請求項13】 前記認識工程の結果に基づいて、前記返答を受け付ける前記ネットワーク上のサーバに適応した形式の返答データを作成する工程と、

前記返答データを前記サーバに提供する工程と、を含むことを特徴とする請求項12に記載のデータ処理方法。

【請求項14】 前記返答が、予め定められた複数の項目から選択すべきものである場合に、

各々の前記項目に係る音声データを認識するための認識文法を作成する工程と

前記認識文法に基づいて、前記ブラウザ装置からの前記返答に係る音声データが、いずれの前記項目に相当するかを判定する工程と、

各々の前記項目に対応して前記返答を受け付ける前記ネットワーク上のサーバに適応した形式の返答データを作成する工程と、

前記返答データを前記サーバに提供する工程と、を含むことを特徴とする請求項11に記載のデータ処理方法。

【請求項15】 前記返答データは、前記識別子を付加したデータを前記ブラウザ装置に対して提供する前に作成されることを特徴とする請求項14に記載のデータ処理方法。

【請求項16】 ネットワーク上で提供されるデータの内容を、ブラウザ装

置に対して音声データとして提供するために、コンピュータを、

前記データに基づいて、その一部又は全部の内容を示す音声データを作成する手段、

作成した音声データを格納する手段、

前記データに、前記音声データを格納した場所を示す識別子を付加したデータを作成する手段、

前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段、  
として機能させるプログラムを記録した記録媒体。

【請求項17】 ネットワーク上で提供されるデータに対して、ブラウザ装置から音声により返答するために、コンピュータを、

前記データの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する手段、

前記データに、前記ブラウザ装置が音声データにより前記返答を行う場合の受付先を示す識別子を付加したデータを作成する手段、

前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段、  
として機能させるプログラムを記録した記録媒体。

【請求項18】 前記プログラムが、コンピュータを、

前記返答に係る音声データが、前記ブラウザ装置から前記受付先へ供給された場合に、前記音声データに対して、音声認識を行う認識手段、  
として機能させるプログラムを含むことを特徴とする請求項17に記載の記録媒体。

【請求項19】 前記プログラムが、コンピュータを、

前記認識手段の結果に基づいて、前記返答を受け付ける前記ネットワーク上のサーバに適応した形式の返答データを作成する手段、

前記返答データを前記サーバに提供する手段、  
として機能させるプログラムを含むことを特徴とする請求項18に記載の記録媒体。

【請求項20】 前記プログラムが、前記返答が、予め定められた複数の項目から選択すべきものである場合に、コンピュータを、

各々の前記項目に係る音声データを認識するための認識文法を作成する手段、  
前記認識文法に基づいて、前記ブラウザ装置からの前記返答に係る音声データ  
が、いずれの前記項目に相当するかを判定する手段、

各々の前記項目に対応して前記返答を受け付ける前記ネットワーク上のサーバ  
に適応した形式の返答データを作成する手段、

前記返答データを前記サーバに提供する手段、  
として機能させるプログラムを含むことを特徴とする請求項 17 に記載の記録媒  
体。

【請求項 21】 前記返答データは、前記識別子を付加したデータを前記ブ  
라우저装置に対して提供する前に作成されることを特徴とする請求項 20 に記載  
の記録媒体。

【請求項 22】 前記データが、マークアップ言語で記述されたデータであ  
り、前記識別子は、該マークアップ言語に対応したタグとして前記データに付加  
されることを特徴とする請求項 1 乃至 6 のいずれかに記載のデータ処理装置。

【請求項 23】 前記データが、マークアップ言語で記述されたデータであ  
り、前記識別子は、該マークアップ言語に対応したタグとして前記データに付加  
されることを特徴とする請求項 7 乃至 9 のいずれかに記載のブラウザシステム。

【請求項 24】 前記データが、マークアップ言語で記述されたデータであ  
り、前記識別子は、該マークアップ言語に対応したタグとして前記データに付加  
されることを特徴とする請求項 10 乃至 15 のいずれかに記載のデータ処理方法  
。

【請求項 25】 前記データが、マークアップ言語で記述されたデータであ  
り、前記識別子は、該マークアップ言語に対応したタグとして前記データに付加  
されることを特徴とする請求項 16 乃至 21 のいずれかに記載の記録媒体。

【請求項 26】 音声を入力する手段と、  
入力された音声に基づいて音声データを作成する手段と、  
作成した音声データを与えられた識別子が示す受付先に供給する手段と、  
を備えたことを特徴とするブラウザ装置。

【請求項 27】 与えられた第 2 の識別子に示された場所から音声データを

取得し、該音声データに係る音声を出力する手段を備えたことを特徴とする請求項 2 6 に記載のブラウザ装置。

【請求項 2 8】 サーバ及びブラウザ装置とネットワークを介して通信可能なデータ処理装置であって、

前記サーバより提供されるデータに基づいて、その一部又は全部の内容を示す音声データを作成する手段と、

作成した音声データを格納する手段と、

前記提供されるデータに、前記音声データを格納した場所を示す第 1 の識別子を付加する手段と、

前記提供されるデータの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する手段と、

前記返答を要求する内容がある場合に、前記第 1 の識別子を付加したデータに返答の受付先を示す第 2 の識別子を更に付加する手段と、

前記第 1 の識別子、或いは、前記第 1 及び第 2 の識別子、を付加したデータを前記ブラウザ装置に対して提供する手段と、

前記ブラウザ装置から前記受付先へ前記返答に係る音声データが供給された場合に、当該音声データに対して音声認識を行う認識手段と、

前記認識手段の認識結果に基づいて、前記返答を受け付ける前記サーバに適応した形式の返答データを作成する手段と、

前記返答データを前記サーバに提供する手段と、  
を備えたことを特徴とするデータ処理装置。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、ネットワークを介してサーバとクライアントとの間で行われる情報の入出力を、クライアント側において音声により実現するブラウザシステム等に関する。

【0 0 0 2】

【従来の技術】

音声入出力機能を有する従来のボイスブラウザシステムとしては、例えば特開平10-124293で提案されている音声司令可能なコンピュータのように、クライアント側で音声合成と音声認識を行なうものがある。しかし、この構成のボイスブラウザシステムでは、クライアントが携帯端末のような計算資源の少ないハードウェアで実現される場合に、クライアントでの処理負荷が資源に対して過大すぎるという問題があった。

【0003】

そこで、例えば、特開平11-110186で提案されているブラウザシステム又は音声プロキシサーバのように、クライアントを実現するハードウェアとは異なるハードウェアで音声合成や認識を行なうボイスブラウザシステムが考案されている。

【0004】

【発明が解決しようとする課題】

しかしながら、上記従来例のボイスブラウザシステムでは、HTML言語等のマークアップ言語により記述されたデータを表示するブラウザプロセスと、音声合成や音声認識による音声の出力や入力を行なうプロセスとが別個であるため、音声合成や音声認識を行なうハードウェアとクライアントを実現するハードウェアとの間で、HTTP等により実現されるHTML言語等のデータ送受の通信とは別に、音声出力および音声入力データの送受の通信を行なわなければならない。

【0005】

そのため、複雑な通信制御や各プロセス間の同期を取る制御などが必要となり、ボイスブラウザシステムの構築が難しくなるという問題があった。さらに、クライアントとサーバとの間で、HTTP以外の通信を許さないファイアウォールが設けられることも多く、こうした場合に他の通信ができずボイスブラウザシステムの構築が困難になるという問題点もあった。

【0006】

従って、本発明の目的は、ネットワーク上で提供されるデータの表示と、該データに対する音声の出力又は入力と、を共通の通信処理の中で処理し得るデータ

処理装置及び方法、ブラウザシステム、ブラウザ装置、記録媒体を提供することにある。

【0007】

【課題を解決するための手段】

本発明によれば、ネットワーク上で提供されるデータの内容を、ブラウザ装置に対して音声データとして提供するためのデータ処理装置であって、前記データに基づいて、その一部又は全部の内容を示す音声データを作成する手段と、作成した音声データを格納する手段と、前記データに、前記音声データを格納した場所を示す識別子を付加したデータを作成する手段と、前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、を備えたことを特徴とするデータ処理装置が提供される。

【0008】

また、本発明によれば、ネットワーク上で提供されるデータに対して、ブラウザ装置から音声により返答するためのデータ処理装置であって、前記データの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する手段と、前記データに、前記ブラウザ装置が音声データにより前記返答を行う場合の受付先を示す識別子を付加したデータを作成する手段と、前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、を備えたことを特徴とするデータ処理装置が提供される。

【0009】

また、本発明によれば、ブラウザ装置と、ネットワークを介して前記ブラウザ装置にデータを提供するサーバと、前記サーバにより提供されるデータの内容を、ブラウザ装置に対して音声データとして提供するためのデータ処理装置と、を備えたブラウザシステムであって、

前記データ処理装置は、前記サーバが提供する前記データに基づいて、その一部又は全部の内容を示す音声データを作成する手段と、作成した音声データを格納する手段と、前記データに、前記音声データを格納した場所を示す識別子を付加したデータを作成する手段と、前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、を備え、



前記ブラウザ装置は、前記識別子に示された場所から前記音声データを取得し、該音声データに係る音声を出力する手段を備えたことを特徴とするブラウザシステムが提供される。

## 【 0 0 1 0 】

また、本発明によれば、ブラウザ装置と、ネットワークを介して前記ブラウザ装置にデータを提供するサーバと、前記サーバにより提供されるデータに対して、ブラウザ装置から音声により返答するためのデータ処理装置と、を備えたブラウザシステムであって、

前記データ処理装置が、前記データの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する手段と、前記データに、前記ブラウザ装置が音声データにより前記返答を行う場合の受付先を示す識別子を付加したデータを作成する手段と、前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、前記返答に係る音声データが、前記ブラウザ装置から前記受付先へ供給された場合に、前記音声データに対して、音声認識を行う認識手段と、前記認識手段の結果に基づいて、前記返答を受け付ける前記サーバに適応した形式の返答データを作成する手段と、前記返答データを前記サーバに提供する手段と、を備え、

前記ブラウザ装置が、音声を入力する手段と、入力された音声に基づいて音声データを作成する手段と、作成した音声データを前記識別子が示す受付先へ供給する手段と、を備えたことを特徴とするブラウザシステムが提供される。

## 【 0 0 1 1 】

また、本発明によれば、ブラウザ装置と、ネットワークを介して前記ブラウザ装置にデータを提供するサーバと、前記サーバにより提供されるデータの内容を、ブラウザ装置に対して音声データとして提供すると共に、前記サーバにより提供されるデータに対して、前記ブラウザ装置から音声により返答するためのデータ処理装置と、を備えたブラウザシステムであって、

前記データ処理装置は、前記サーバが提供する前記データに基づいて、その一部又は全部の内容を示す音声データを作成する手段と、作成した音声データを格納する手段と、前記データに、前記音声データを格納した場所を示す第 1 の識別

子を付加したデータを作成する手段と、前記第 1 の識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、前記データの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する手段と、前記データに、前記ブラウザ装置が音声データにより前記返答を行う場合の受付先を示す第 2 の識別子を付加したデータを作成する手段と、前記第 2 の識別子を付加したデータを前記ブラウザ装置に対して提供する手段と、前記返答に係る音声データが、前記ブラウザ装置から前記受付先へ供給された場合に、前記音声データに対して、音声認識を行う認識手段と、前記認識手段の結果に基づいて、前記返答を受け付ける前記サーバに適応した形式の返答データを作成する手段と、前記返答データを前記サーバに提供する手段と、を備え、

前記ブラウザ装置は、前記第 1 の識別子に示された場所から前記音声データを取得し、該音声データに係る音声を出力する手段と、音声を入力する手段と、

入力された音声に基づいて音声データを作成する手段と、作成した音声データを前記第 2 の識別子が示す受付先へ供給する手段と、を備えたことを特徴とするブラウザシステムが提供される。

#### 【 0 0 1 2 】

また、本発明によれば、ネットワーク上で提供されるデータの内容を、ブラウザ装置に対して音声データとして提供するためのデータ処理方法であって、前記データに基づいて、その一部又は全部の内容を示す音声データを作成する工程と、作成した音声データを格納する工程と、前記データに、前記音声データを格納した場所を示す識別子を付加したデータを作成する工程と、前記識別子を付加したデータを前記ブラウザ装置に対して提供する工程と、を含むことを特徴とするデータ処理方法が提供される。

#### 【 0 0 1 3 】

また、本発明によれば、ネットワーク上で提供されるデータに対して、ブラウザ装置から音声により返答するためのデータ処理方法であって、前記データの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する工程と、前記データに、前記ブラウザ装置が音声データにより前記返答を行う場合の受付先を示す識別子を付加したデータを作成する工程と、前記識別子を付加した

データを前記ブラウザ装置に対して提供する工程と、を含むことを特徴とするデータ処理方法が提供される。

【 0 0 1 4 】

また、本発明によれば、ネットワーク上で提供されるデータの内容を、ブラウザ装置に対して音声データとして提供するために、コンピュータを、前記データに基づいて、その一部又は全部の内容を示す音声データを作成する手段、作成した音声データを格納する手段、前記データに、前記音声データを格納した場所を示す識別子を付加したデータを作成する手段、前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段、として機能させるプログラムを記録した記録媒体が提供される。

【 0 0 1 5 】

また、本発明によれば、ネットワーク上で提供されるデータに対して、ブラウザ装置から音声により返答するために、コンピュータを、前記データの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する手段、前記データに、前記ブラウザ装置が音声データにより前記返答を行う場合の受付先を示す識別子を付加したデータを作成する手段、前記識別子を付加したデータを前記ブラウザ装置に対して提供する手段、として機能させるプログラムを記録した記録媒体が提供される。

【 0 0 1 6 】

また、本発明によれば、音声を入力する手段と、入力された音声に基づいて音声データを作成する手段と、作成した音声データを与えられた識別子を示す受付先に供給する手段と、を備えたことを特徴とするブラウザ装置が提供される。

【 0 0 1 7 】

また、本発明によれば、サーバ及びブラウザ装置とネットワークを介して通信可能なデータ処理装置であって、前記サーバより提供されるデータに基づいて、その一部又は全部の内容を示す音声データを作成する手段と、作成した音声データを格納する手段と、前記提供されるデータに、前記音声データを格納した場所を示す第 1 の識別子を付加する手段と、前記提供されるデータの内容に、前記ブラウザ装置からの返答を要求する内容があるか否かを判定する手段と、前記返答

を要求する場合がある場合に、前記第 1 の識別子を付加したデータに返答の受付先を示す第 2 の識別子を更に付加する手段と、前記第 1 の識別子、或いは、前記第 1 及び第 2 の識別子、を付加したデータを前記ブラウザ装置に対して提供する手段と、前記ブラウザ装置から前記受付先へ前記返答に係る音声データが供給された場合に、当該音声データに対して音声認識を行う認識手段と、前記認識手段の認識結果に基づいて、前記返答を受け付ける前記サーバに適応した形式の返答データを作成する手段と、前記返答データを前記サーバに提供する手段と、を備えたことを特徴とするデータ処理装置が提供される。

【0018】

【発明の実施の形態】

以下、図面を参照して本発明の好適な実施の形態について説明する。

【0019】

図 1 は、本発明の一実施形態に係るボイスブラウザシステムの構成図である。

【0020】

ボイスブラウザシステムは、ネットワーク 104 により相互に接続された、ブラウザ装置としてのクライアントコンピュータ 101 と、クライアントコンピュータ 101 に対して情報を提供するサーバコンピュータ 103 と、サーバコンピュータ 103 により提供される情報を音声データとしてクライアントコンピュータ 101 に提供し、また、クライアントコンピュータ 101 からサーバコンピュータ 103 に対して行われる要求又は返答を音声データにより実現するボイスゲートウェイコンピュータ 102 と、により構成される。

【0021】

クライアントコンピュータ 101、ボイスゲートウェイコンピュータ 102、サーバコンピュータ 103 は、いずれも通信デバイスを備えており、ネットワーク 104 を介して TCP/IP プロトコルにより実現される HTTP を互いに送受信できる。

【0022】

ここで、サーバコンピュータ 103 としては、従来の Web サーバと呼ばれるコンピュータ、すなわち HTTP リクエストに対して、その URL に合致するデ

ータをHTTPレスポンスとして返すWebサーバプログラムを実行するコンピュータを採用することができる。

【0023】

また、ネットワーク104としては、例えば、インターネットあるいはイントラネットを挙げることができる。

【0024】

このボイスブラウザシステムにおいて、クライアントコンピュータ101とサーバコンピュータ103との間における通信は、ボイスゲートウェイコンピュータ102を介して行うことができる。

【0025】

この場合、クライアントコンピュータ101からサーバコンピュータ103への要求、返答等をHTTPリクエストといい、特に、クライアントコンピュータ101からボイスゲートウェイコンピュータ102への要求、返信等をクライアント側HTTPリクエストとも称し、これに対応するボイスゲートウェイコンピュータ102からサーバコンピュータ103への要求、返答等をサーバ側HTTPリクエストとも称す。

【0026】

また、HTTPリクエストに対するサーバコンピュータ103からクライアントコンピュータ103への情報の提供等をHTTPレスポンスといい、特に、ボイスゲートウェイコンピュータ102に対する返答をサーバ側HTTPレスポンスとも称し、これに対応するボイスゲートウェイコンピュータ102からクライアントコンピュータ101に対する返答をクライアント側HTTPレスポンスとも称す。

【0027】

次に、クライアントコンピュータ101とボイスゲートウェイコンピュータ102について詳細に説明する。

【0028】

クライアントコンピュータ101は、ディスプレイ、キーボードやマウスなどの入力デバイス、及び、スピーカやマイクロホン等の音声入出力デバイス、を備

えており、サーバコンピュータ103により提供されるHTML言語で記述されたデータの表示、及び、該データの内容を示すボイスゲートウェイコンピュータ102からの音声データの音声出力、並びに、音声データを含むクライアント側HTTPリクエストの作成、若しくは、文字入力を可能とするコンピュータである。

【0029】

図2は、ブラウザ装置として機能するクライアントコンピュータ101の基本構成を示すブロック図である。

【0030】

図2において、201は、HTTPレスポンスを受信するレスポンス受信部である。202は、レスポンス受信部で受信されたHTTPレスポンスに含まれるHTMLデータを解析するHTMLデータ解析部である。203は、HTMLデータ解析部202で解析された結果に従ってデータを表示するHTML表示部である。

【0031】

204は、HTMLデータ解析部202で解析された結果に従って音声データが含まれる場合はその音声出力を行なう音声出力部である。205は、ユーザからのキーボード等の入力デバイスからの入力を受け付ける直接入力部である。207は、音声入力を受け付ける音声入力部である。

【0032】

206は、サーバコンピュータ103から提供され、HTML表示部203で表示される情報に対する指示操作を受け付ける指示要求入力部である。208は、直接入力部205への入力、指示要求入力部206への入力、音声入力部207への入力のいずれかあるいはその組み合わせからHTTPリクエストを作成するリクエスト作成部である。209は、リクエスト作成部で作成されたHTTPリクエストを送信するリクエスト送信部である。

【0033】

一方、ボイスゲートウェイコンピュータ102は、クライアントコンピュータ101から送信されたクライアント側HTTPリクエストに音声データが含まれ

る場合はそれを音声認識し、認識結果に基づいてサーバ側HTTPリクエストを作成して送信し、また、HTTPリクエストに対するサーバコンピュータ103によるサーバ側HTTPレスポンスに含まれるHTMLデータから、クライアントコンピュータ101で音声出力すべき音声データを作成し、HTMLデータと共にその音声データをクライアントコンピュータ101へ提供するデータ処理装置として機能する。

## 【0034】

図5は、ボイスゲートウェイコンピュータ102の基本構成を示すブロック図である。

## 【0035】

図5において、501はクライアントコンピュータ101から送信されるクライアント側HTTPリクエストを受信するリクエスト受信部である。502は、リクエスト受信部501で受信されたクライアント側HTTPリクエストに音声データが含まれる場合に、それを音声認識する音声認識部である。

## 【0036】

503は、音声認識部502の認識結果に基づいて、音声データを含むクライアント側HTTPリクエストを、サーバコンピュータ103に適応した形式のサーバ側HTTPリクエストにデータ変換するリクエスト変換部である。

## 【0037】

504は、サーバ側HTTPリクエストをサーバコンピュータ103に送信するリクエスト送信部であり、対応するクライアント側HTTPリクエストが音声データを含む場合は、リクエスト変換部503で変換されたものを、音声データを含まない場合は、リクエスト受信部501で受信されたものを、それぞれサーバコンピュータ103に送信する。

## 【0038】

505は、リクエスト送信部504が送信したサーバ側HTTPリクエストに対するサーバ側HTTPレスポンスを受信するレスポンス受信部である。506は、レスポンス受信部505で受信されたサーバ側HTTPレスポンスに含まれるHTMLデータを解析するHTMLデータ解析部である。

## 【0039】

507は、HTMLデータ解析部506で解析された結果を用いて音声出力すべき音声データを作成する音声合成部である。508は、HTMLデータ解析部506で解析された結果を用いて、次に受け付ける音声入力候補とその入力があった場合に作成すべきリクエストを作成する入力候補作成部である。すなわち、入力候補作成部508は、サーバコンピュータ103がクライアントコンピュータ101に提供した情報の中に、クライアントコンピュータ101からの返答を要求する内容があるか否かを判定し、該内容があればその候補を作成する。

## 【0040】

509は、入力候補作成部508で作成された入力候補から音声認識部502が使用する音声認識文法を作成する認識文法作成部である。510は、認識文法作成部509で作成された認識文法を保持する認識文法保持部である。

## 【0041】

511は、入力候補作成部508で作成された入力候補とその入力候補に対応する音声入力があったときに作成すべきHTTPリクエストの対を保持する次リクエスト保持部である。

## 【0042】

512は、音声合成部507で作成された音声データを保持する音声データ保持部である。513は、音声データ保持部512に保持されている音声データを取得するURLと、次の音声認識処理を起動するURLとを、レスポンス受信部505で受信されたレスポンスデータに含まれるHTMLデータに埋め込むHTMLデータ変換部である。

## 【0043】

514は、HTMLデータ変換部で作成されたHTMLデータをクライアントコンピュータ101へのクライアント側HTTPレスポンスとして返すレスポンス送信部である。515は、リクエスト受信部501で受信したクライアント側HTTPリクエストが、音声データ保持部512に保持されている音声データの要求であった場合に、その音声データをクライアントコンピュータ101へのクライアント側HTTPレスポンスとして返す音声データ返信部である。



【0044】

次に、図3は、クライアント側HTTPレスポンスに含まれるHTMLデータの例である。本実施形態では、マークアップ言語としてHTML4.0として定義されているHTML言語仕様に加えて、VOICEOUT、VOICEINの二つの拡張タグを使用している。

【0045】

ここで、VOICEOUTは、音声合成部507で作成された音声データの格納場所、すなわち、音声データ保持部512を示す識別子としてのタグである。

【0046】

また、VOICEINは、クライアントコンピュータ101からのHTTPリクエストが音声データを含む場合に、その受付先、リクエスト受信部501を示す識別子としてのタグである。

【0047】

VOICEOUTが出現した場合、クライアントコンピュータ101は、その属性で指定されるURLにある音声データを、別のHTTPリクエストを発行して取得し、取得した音声データをスピーカ等音声出力デバイスから出力する。

【0048】

また、VOICEINタグは、クライアントコンピュータ101のマイクロホン等の音声入力デバイスに音声入力があった場合に作成するHTTPリクエストのベースとなるベースURLを指定する。

【0049】

この例を用いて、クライアントコンピュータ101における処理を説明する。

【0050】

レスポンス受信部201は、何らかのHTTPリクエストに対するHTTPレスポンスとして、図3に示したHTMLデータを含むHTTPレスポンスを受信する。受信の手順は、従来のブラウザ装置と同様である。

【0051】

HTMLデータ解析部202は、一般的なHTMLデータの解析を行うが、特に、VOICEOUT、VOICEINタグについても動作に必要なデータを抽出する点だけが

異なる。HTML表示部203は、従来のブラウザ装置と同様にHTMLデータに基づき情報を表示する。

【0052】

音声出力部204は、VOICEOUTタグが解析された場合に、その属性で示されるURLで示す音声データを要求するHTTPリクエストを送信し、ボイスゲートコンピュータ103からの対応するHTTPレスポンスのボディに含まれる音声データを取得する。

【0053】

そして、取得した音声データをスピーカー等の音声出力デバイスにより出力する。

【0054】

一方、音声入力部207は、マイクロホン等の音声入力デバイスに入力があると、それをA/D変換して例えばPCMデータ等の音声データを作成する。この場合、PCMデータの始点と終点を定めるタイミングは、音声入力のパワーが閾値を越えている間を取るとか、何らかのキーが押下されている間を取るといった方法で定められる。

【0055】

リクエスト作成部208は、音声入力部207に入力があった場合は、HTMLデータ解析部202が解析した結果のうちで、VOICEINタグの属性として示されるURLへのPOSTリクエストを作成する。そして、そのボディに音声入力部207が作成したPCMデータを入れる。図3の例で示すHTMLデータを出力している状態のクライアントコンピュータ101に、音声入力があった場合のHTTPリクエストは図4のように作成される。リクエスト送信部209は、リクエスト作成部208で作成されたHTTPリクエストを、そのURLが示すコンピュータ、すなわちボイスゲートウェイコンピュータ102へ送信する。

【0056】

次に、図6に示すフローチャートを用いてボイスゲートウェイコンピュータ102における処理の概要を説明する。

【0057】

処理のメインルーチンでは、HTTPを受けるポート（通常80番ポートが使用されるがこの限りではない）への接続要求を待ち受けている。クライアントコンピュータ101から接続要求があると接続を確立する。そして、ここで説明する処理に移る。本実施形態では、説明を簡単にするため、待ち受け処理と同一の単一スレッドで処理を実現する場合で説明するが、マルチスレッドで実現してもよい。なお、本処理が終了すると、再びポートの接続要求待ち処理に戻る。

## 【0058】

まず、ステップS601では、クライアントコンピュータ101からのHTTPリクエストを受信する。そして、ステップS602に移る。

## 【0059】

ステップS602では、HTTPリクエストデータ中のURLを取り出し、それがボイスゲートウェイコンピュータ102の“/out.wav”を指すものであればステップS613に移る。指すものでなければステップS603に移る。

## 【0060】

ステップS603では、前記URLがボイスゲートウェイコンピュータ102の“/voicein”を指すものであればステップS604に移る。指すものでなければステップS606に移る。

## 【0061】

ステップS604では、HTTPリクエストのボディを取り出し、それを音声データとして音声認識を行なう。この時に認識文法保持部510に保持されている認識文法が使用される。そしてステップS605に移る。

## 【0062】

ステップS605では、ステップS604で認識された結果に対する次HTTPリクエストを次リクエスト保持部511から取り出す。そしてステップS607に移る。

## 【0063】

ステップS606では、クライアントコンピュータ101から送られてきたHTTPリクエストのURLをそのまま次HTTPリクエストとする。そしてステップS607に移る。

## 【0064】

ステップS607では、次HTTPリクエストのURL中のホスト部が示すサーバ（サーバコンピュータ103）に次HTTPリクエストを送信し、HTTPレスポンスを得る。この動作は、従来のプロキシ装置と同じである。そして、ステップS608に移る。なお、図7は、HTTPレスポンスのデータの例を示す図である。

## 【0065】

ステップS608では、ステップS607で受領したHTTPレスポンス（サーバ側HTTPレスポンス）のボディにあるHTMLデータを解析する。解析することで、HTMLデータ中の各タグの木構造とタグの要素を取り出すことができるようになる。そして、ステップS609に移る。

## 【0066】

ステップS609では、ステップS608の解析結果を用いて、クライアントコンピュータ101で音声出力すべき音声データを作成する。HTMLデータ中の一部あるいは全部のテキストを音声合成して音声データを作成する。どのテキストを音声合成するかは任意に定めることができ、ここでは簡単に最初のPタグ要素を音声合成するとして説明を続ける。図7のデータ例では、「製品の種類を選択して下さい。」という合成音声データが作成される。合成音声データは、WAVEフォーマットのファイルとして、“/out.wav”で参照される場所に格納する。そして、ステップS610に移る。

## 【0067】

ステップS610では、当該HTMLデータと音声データとをクライアントコンピュータ101へ出力した後、その内容によりクライアントコンピュータ101から受け付ける返答の音声入力候補を作成する。どのような音声入力を受理するかは、任意に定めることができるが、ここでは簡単にSELECTタグ中の各OPTIONタグの要素を入力候補とする。図7の例では、「コピー」、「プリンタ」、「Fax」が入力候補となる。

## 【0068】

そして、入力候補の各単語を認識する認識文法を作成する。さらに、各要素が

音声入力された時の次HTTPリクエストとして、対応するSELECTタグが選択されてフォームがサブMITTされた時に生成されるリクエストURLを作成し、それを次リクエスト保持部511に保持する。

【0069】

図8は、図7のデータを処理した場合の、次リクエスト保持部511のデータ構成例を示した図である。同図において、各行が一つの入力候補に対応する。列801には、入力候補の文字列が保持される。列802には、次HTTPリクエストのURLが保持される。そして、ステップS611に移る。

【0070】

ステップS611では、当該HTMLデータ中に、VOICEOUTタグとVOICEINタグとを埋め込む。本実施形態では、それぞれのURLは固定であるので、常に同じタグパターンが埋め込まれることとなる。そして、タグを埋め込んだHTMLデータをクライアント側レスポンスとして、ステップS613に移る。

【0071】

ステップS612では、一つ前の処理のステップS609で格納した音声データに係るクライアント側レスポンスを作成し、ステップS613に移る。

【0072】

ステップS613では、作成したクライアント側HTTPレスポンスをクライアントコンピュータ101に提供する。そして、クライアントコンピュータ101との接続を切断し処理を終了する。

【0073】

最後に、図9を用いて本実施形態における各コンピュータ間での通信例を示す。

【0074】

始めに、クライアントコンピュータ101（ブラウザ）にURLが直接入力され、ボイスゲートウェイコンピュータ102にhttp://server/index.htmlを要求するクライアント側HTTPリクエストが送られる（901）。なお、URLの直接入力ではなく、同URLをアンカーに持つオブジェクトをブラウザ表示画面上で指示することで、そのURLを要求するHTTPリクエストが送られる場合

もある。この送信については、従来のブラウザ装置と同様である。

【0075】

次に、ボイスゲートウェイコンピュータ102は、サーバ103のURLに対するHTTPリクエストなので、従来のプロキシ装置と同様に、サーバ103に対し/index.htmlを要求する新たなサーバ側HTTPリクエストを送る(902)。

【0076】

サーバ103は、/index.htmlで指示されるデータをボディに含むサーバ側HTTPレスポンスをボイスゲートウェイコンピュータ102に返す(903)。例えば、図7がこのHTTPレスポンスの例である。

【0077】

ボイスゲートウェイコンピュータ102は、受信したサーバ側HTTPレスポンスをもとに音声データや入力候補などを作成し、新たなタグを埋め込んだHTMLデータ(例えば図4)をボディに含むクライアント側HTTPレスポンスをクライアントコンピュータ101に返す(904)。

【0078】

クライアントコンピュータ101では、受信したクライアント側HTTPレスポンスのボディに入っているHTMLデータを表示すると共に、VOICEOUTタグで指示される音声データ(図4の例ではhttp://gateway/out.wav)を要求するクライアント側HTTPリクエストをボイスゲートウェイコンピュータ102に送る(905)。

【0079】

ボイスゲートウェイコンピュータ102は、out.wavで指示される音声データをクライアントコンピュータ101に返信する(906)。この音声データは、クライアント側HTTPレスポンス(904)を提供する前に作成、格納されたデータである。

【0080】

ここで、クライアントコンピュータ101に音声入力があると、その音声データをボディに含むクライアント側HTTPリクエスト(POSTリクエスト)が

クライアントコンピュータ101からボイスゲートウェイコンピュータ102に送られる(907)。例えば、図3に示したデータが送られる。

【0081】

ボイスゲートウェイコンピュータ102は、受信したPOSTリクエストのボディに含まれる音声データに対して音声認識を行なう。ここで、例えば「コピー」と認識されると、次リクエスト保持部511の内容に従い、この例では、サーバ103に対して、/cgi1?category=copyを要求するサーバ側HTTPリクエストを送る(908)。この時に使用される認識文法や次リクエスト保持部511の内容は、先のレスポンス(904)を提供する前に作成されたものである。

【0082】

サーバ103は、受信したサーバ側HTTPリクエストに従い、CGIプログラムを起動して、ボイスゲートウェイコンピュータ102にサーバ側HTTPレスポンスを返す(909)。

【0083】

ボイスゲートウェイコンピュータ102は、レスポンス903を受信したときと同様に、新たに音声データや認識文法を作成して、クライアント側HTTPレスポンスをクライアントコンピュータ101に返す(910)。

【0084】

このように、本実施形態のボイスブラウザシステムでは、ブラウザ(クライアントコンピュータ101)、ボイスゲートウェイ(ボイスゲートウェイコンピュータ102)、サーバ(サーバ103)しか存在せず、これらの間の通信だけでよいので、サーバから提供されるデータの表示と、該データに対応した音声入出力とを共通の通信処理の中で行うことができ、通信管理が簡単になる。また、全ての通信はHTTPで行うことができるので、一般的にHTTPだけを通すファイアーウォールが間にある場合でも、問題なく通信が行なえる。

【0085】

なお、上記実施形態においては、クライアントコンピュータ、ボイスゲートウェイコンピュータ、サーバコンピュータの三つのコンピューターで、それぞれブラウザ、ボイスゲートウェイ、サーバを実現する場合について説明したが、これ

に限定されるものではなく、例えば一つのコンピュータでボイスゲートウェイとサーバの両方を実現してもよい。

【0086】

また、上記実施形態においては、ネットワークに接続されているブラウザ装置、ボイスゲートウェイ装置、サーバ装置が各々1台である場合について説明したが、これに限定されるものではなく、各々複数台あり、要求するデータに応じて任意の装置を利用してもよい。

【0087】

また、上記実施形態においては、ボイスゲートウェイコンピュータ102において、識別子としてVOICEOUTタグとVOICEINタグとをそれぞれ一つずつ付け加える場合について説明したが、これに限定されるものではなく、複数付けてもよい。また、どちらか一方あるいは双方を付けなくてもよい。VOICEOUTを複数にする場合は、音声データを指示するURLをそれぞれ別のものにすることとなる。また、VOICEINを複数にする場合は、認識文法や次HTTPリクエストデータを複数用意し、VOICEINの属性で指示するURLをそれぞれ別のものにすることとなる。そして、ブラウザからボイスゲートウェイに、それらのURLで音声データを含むリクエストが送られてきた場合は、用意した複数の認識文法や次HTTPリクエストデータをURLで判別して使用する。

【0088】

また、上記実施形態においては、合成音声データをWAVEフォーマットで、入力音声データを生のPCMデータで受け渡す場合について説明したが、これに限定されるものではなく、任意の音声フォーマットを使用してもよい。複数の音声フォーマットを許す場合には、音声フォーマットを示すデータを、タグの属性やHTTPヘッダに記すこととなる。

【0089】

また、上記実施形態において使用したタグ名や属性名は、これらに限定されるものではなく、別の名称を用いてもよい。また、属性で表現したものを別のタグにする、あるいはタグで表現したものをHTTPヘッダで表現するというように、HTTPとHTMLの任意の拡張で実現してもよい。



## 【 0 0 9 0 】

また、データ及びプログラムは、HTMLとHTTPに限定されるものではなく、別のマークアップ言語で記述されるデータや別のプロトコルでもよい。例えば、WAPプロトコルでWMLで記述されたデータに同様の音声の埋め込み、認識を行ってもよい。

## 【 0 0 9 1 】

また、上記実施形態は、前述した機能を実現するソフトウェアプログラムを動作させる他、その機能の全部または一部を実現する論理回路により達成されることは言うまでもない。

## 【 0 0 9 2 】

以上、本発明の好適な実施の形態について説明したが、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体（または記録媒体）を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム(OS)などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

## 【 0 0 9 3 】

さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【 0 0 9 4 】

【発明の効果】

以上説明したように、ネットワーク上で提供されるデータの表示と、該データに対する音声の出力又は入力と、を共通の通信処理の中で処理することができる。

【図面の簡単な説明】

【図 1】

本発明の一実施形態に係るボイスブラウザシステムの構成図である。

【図 2】

クライアントコンピュータ 1 0 1 の基本構成を示すブロック図である。

【図 3】

クライアント側 H T T P レスポンスデータの例を示す図である。

【図 4】

クライアント側 H T T P リクエストデータの例を示す図である。

【図 5】

ボイスゲートウェイコンピュータ 1 0 2 の基本構成を示すブロック図である。

【図 6】

ボイスゲートウェイコンピュータ 1 0 2 における処理を示すフローチャートである。

【図 7】

H T T P レスポンスのデータの例を示す図である。

【図 8】

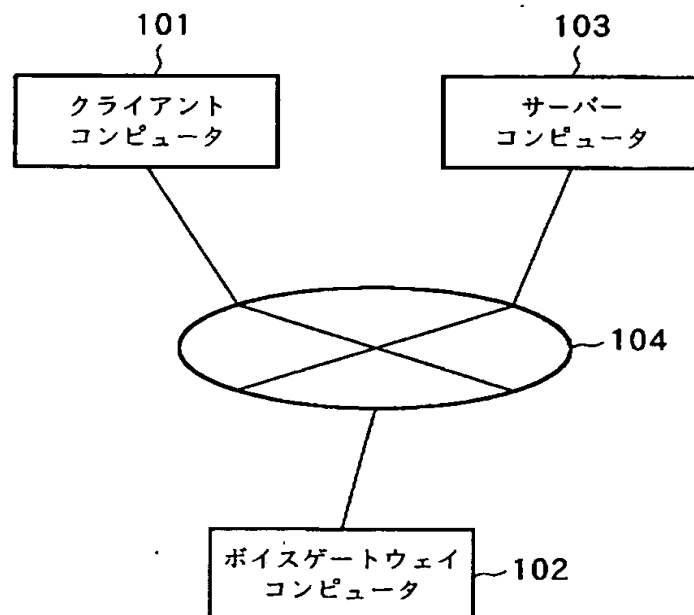
図 7 のデータを処理した場合の、次リクエスト保持部 5 1 1 のデータ構成例を示した図である。

【図 9】

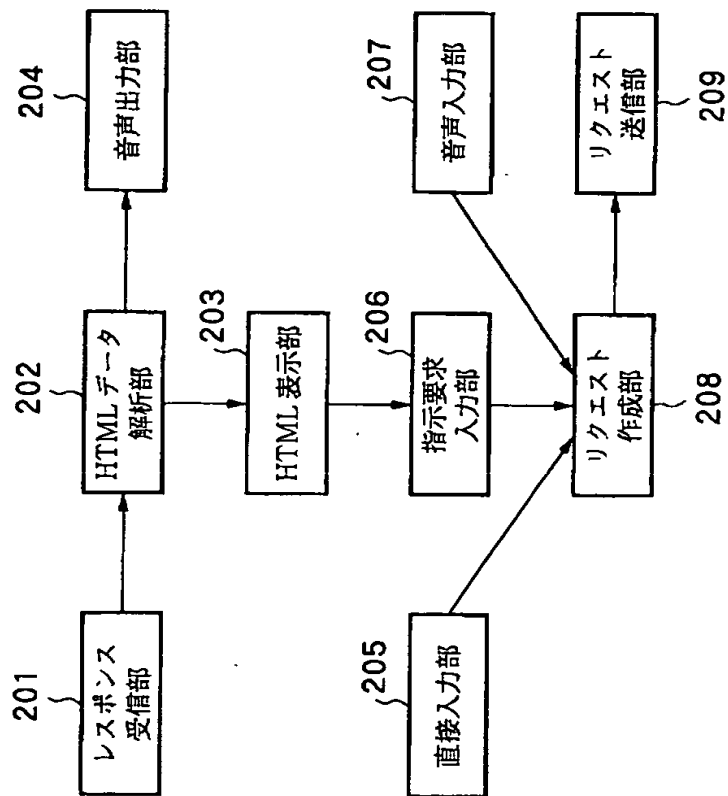
本発明の実施形態における各コンピュータ間の例を示す図である。

【書類名】 図面

【図 1】



【図 2】



【図 3】

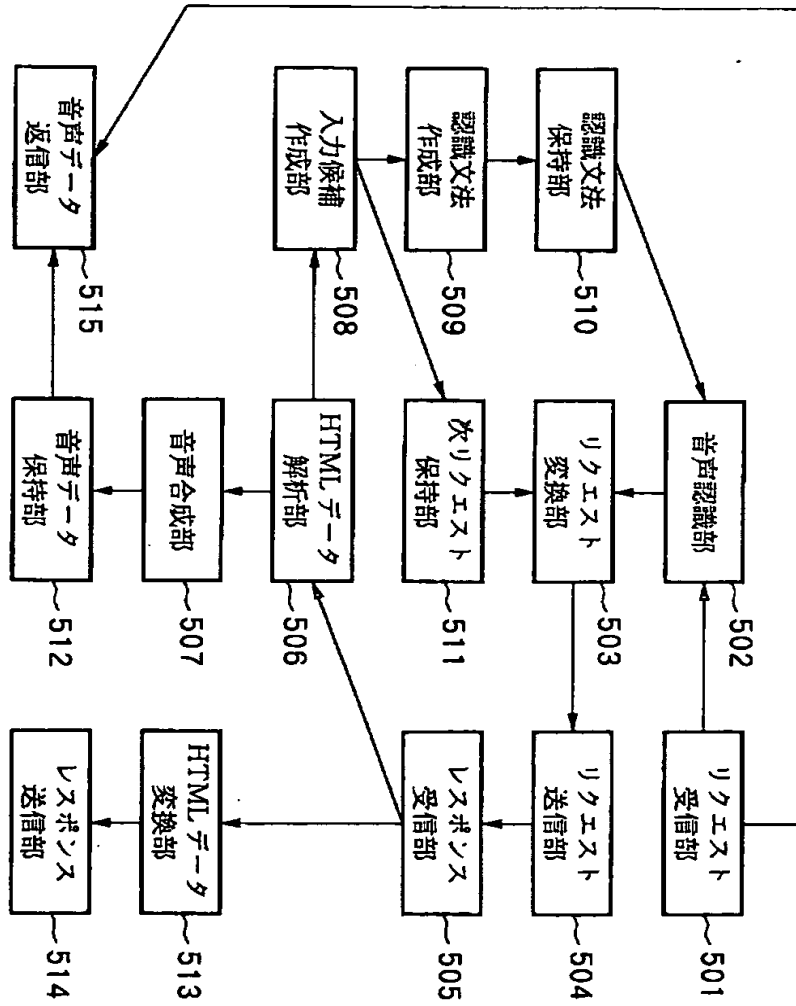
```
POST http://gateway/voicein HTTP/1.0
(PCM データ)
```

【図 4】

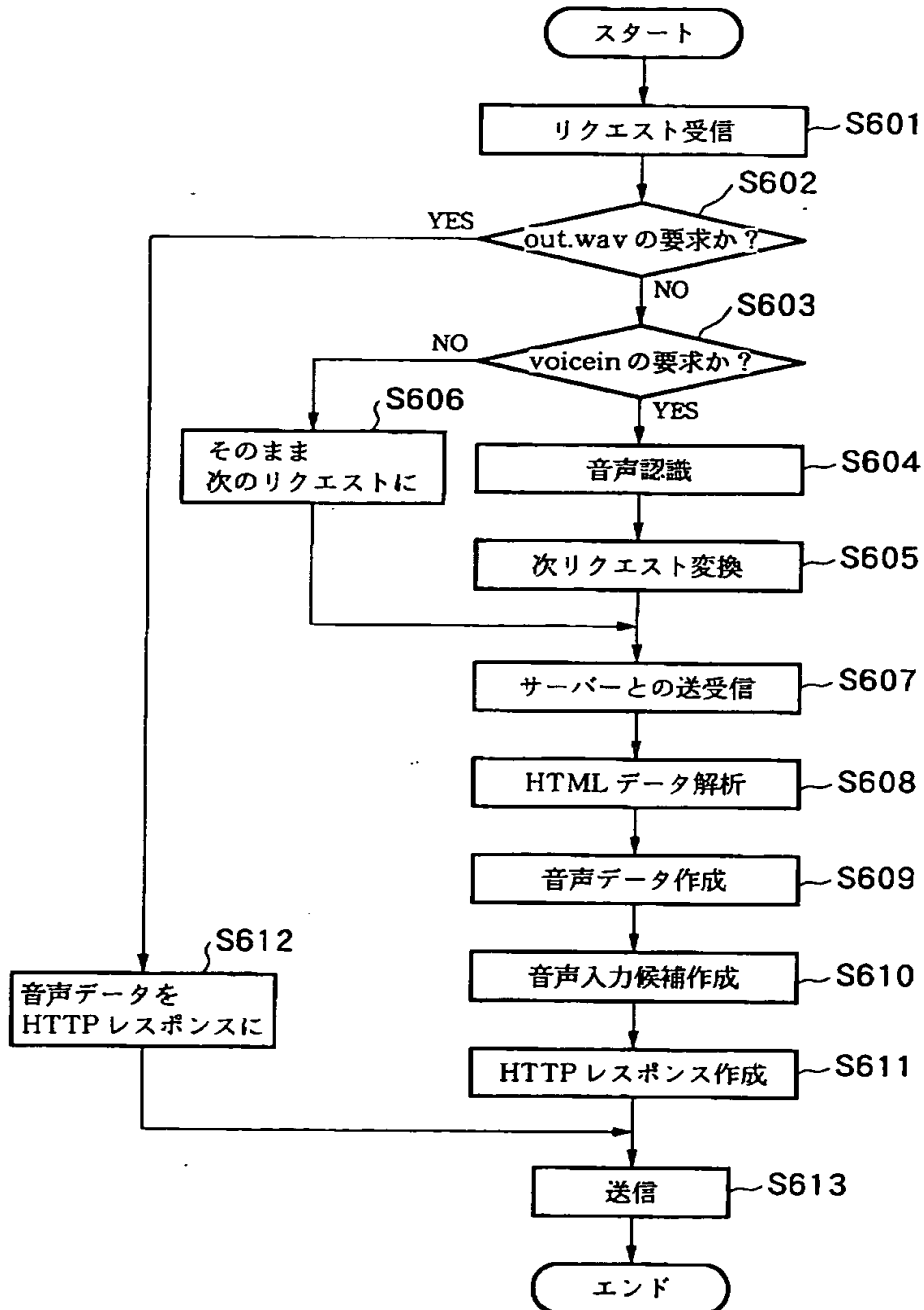
```
HTTP/1.0 200 OK
Content-type:text/html

<HTML> <HEAD> <TITLE>Top</TITLE> </HEAD> <BODY>
<P> 製品の種類を選択して下さい.</P>
<FORM action="/cgil"method="get">
  製品種別< SELECT name="category">
    <OPTION value="copy"> コピー
    <OPTION value="printer"selected> プリンタ
    <OPTION value="fax"> Fax
  </SELECT> <BR>
  <INPUT type="submit" value="表示">
</FORM>
<VOICEOUT href="http://gateway/out.wav">
<VOICEIN action="http://gateway/voicein">
</BODY> </HTML>
```

【図 5】



【図 6】



【図 7】

```

HTTP/1.0 200 OK
Content-type: text/html

<HTML><HEAD><TITLE>Top</TITLE></HEAD><BODY>
<P>製品の種類を選択して下さい.</P>
<FORM action="/cgil"method="get">
  製品種別< SELECT name="category">
    <OPTION value="copy">コピー
    <OPTION value="printer"selected>プリンタ
    <OPTION value="fax">Fax
  </SELECT><BR>
  <INPUT type="submit" value="表示">
</FORM>
</BODY></HTML>

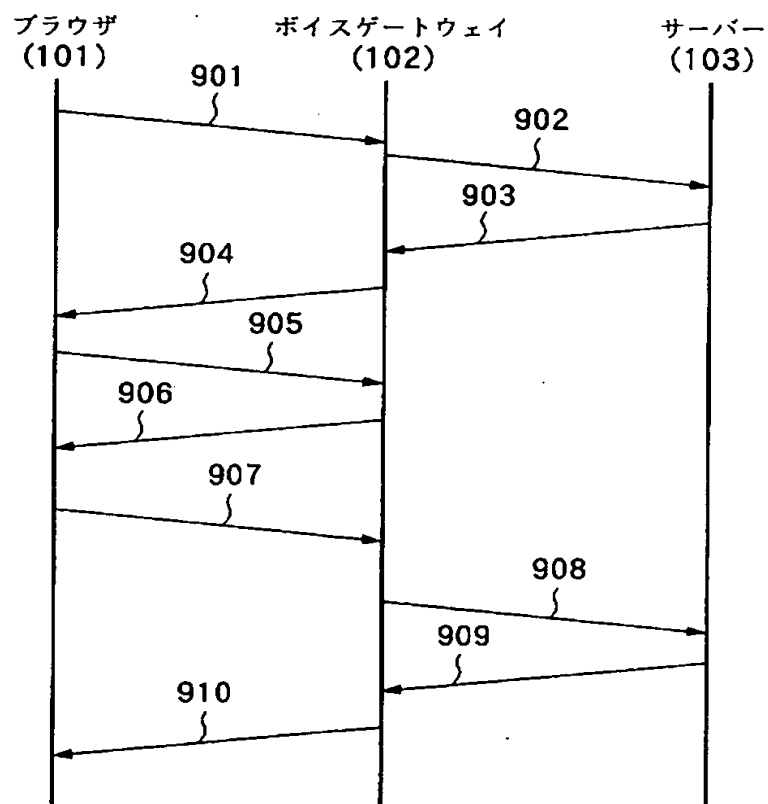
```

【図 8】

801	802
コピー	http://server/cgil?category=copy
プリンタ	http://server/cgil?category=printer
Fax	http://server/cgil?category=fax



【図9】



【書類名】 要約書

【要約】

【課題】 ボイスブラウザシステムにおいて、ネットワーク上で提供されるデータの表示と、該データに対する音声の出力又は入力と、を共通の通信処理の中で処理すること。

【解決手段】 ネットワーク上で提供されるデータの内容を、ブラウザ装置に対して音声データとして提供するために、ゲートウェイ上において、前記データに基づいて、その一部又は全部の内容を示す音声データを作成して格納し、前記データに、音声データを格納した場所を示す識別子<VOICEOUT ...>を付加したデータを作成し、これをブラウザ装置に対して提供する。ブラウザ装置は、識別子の示す場所から音声データを受け取ることとなる。

【選択図】 図 4

出 願 人 履 歴 情 報

識別番号 [000001007]

1. 変更年月日	1990年 8月30日
[変更理由]	新規登録
住 所	東京都大田区下丸子3丁目30番2号
氏 名	キヤノン株式会社